

Attorney Docket No. 1448.1042

THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Itaru SAKOU

Application No.: 10/629,897

Group Art Unit: TO BE ASSIGNED

Filed: July 30, 2003

Examiner: TO BE ASSIGNED

For: METHOD, PROGRAM AND APPARATUS FOR PREDICTING GENE EXPRESSION

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN
APPLICATION IN ACCORDANCE
WITH THE REQUIREMENTS OF 37 C.F.R. § 1.55**

Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

In accordance with the provisions of 37 C.F.R. § 1.55, the applicant(s) submit(s)
herewith a certified copy of the following foreign application:

Japanese Patent Application No(s). 2002-224075

Filed: July 31, 2002

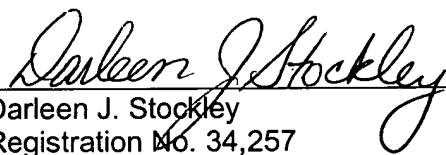
It is respectfully requested that the applicant(s) be given the benefit of the foreign filing
date(s) as evidenced by the certified papers attached hereto, in accordance with the
requirements of 35 U.S.C. § 119.

Respectfully submitted,

STAAS & HALSEY LLP

Date: June 30, 2004

By:


Darleen J. Stockley
Registration No. 34,257

1201 New York Ave, N.W., Suite 700
Washington, D.C. 20005
Telephone: (202) 434-1500
Facsimile: (202) 434-1501

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 2 年 7 月 3 1 日
Date of Application:

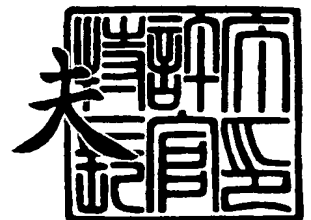
出 願 番 号 特 願 2 0 0 2 - 2 2 4 0 7 5
Application Number:
[ST. 10/C] : [J P 2 0 0 2 - 2 2 4 0 7 5]

出 願 人 富 士 通 株 式 会 社
Applicant(s):

2 0 0 3 年 8 月 6 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



【書類名】 特許願

【整理番号】 0251495

【提出日】 平成14年 7月31日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 17/30

【発明の名称】 遺伝子発現部位予測支援方法、遺伝子発現部位予測支援
プログラムおよび遺伝子発現部位予測支援装置

【請求項の数】 10

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通
株式会社内

 【氏名】 佐甲 格

【特許出願人】

 【識別番号】 000005223

 【氏名又は名称】 富士通株式会社

【代理人】

 【識別番号】 100104190

 【弁理士】

 【氏名又は名称】 酒井 昭徳

【手数料の表示】

 【予納台帳番号】 041759

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9906241

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 遺伝子発現部位予測支援方法、遺伝子発現部位予測支援プログラムおよび遺伝子発現部位予測支援装置

【特許請求の範囲】

【請求項 1】 ゲノム配列上における、発現部位が特定されている周辺遺伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出する算出工程と、

前記算出工程によって算出された距離に基づいて、前記未知の遺伝子の発現部位を特定する特定工程と、

を含んだことを特徴とする遺伝子発現部位予測支援方法。

【請求項 2】 未知の遺伝子の配列情報を入力する入力工程と、

前記入力工程によって入力された配列情報にかかる未知の遺伝子のゲノム配列上の位置を取得する位置情報取得工程と、

前記位置情報取得工程によって取得された未知の遺伝子のゲノム配列上の位置の周辺に存在する、発現部位が特定されている周辺遺伝子を検索する検索工程と、

前記検索工程によって検索された周辺遺伝子によって特定されている発現部位に関する情報を取得する発現部位情報取得工程と、

前記検索工程によって検索された発現部位が特定されている周辺遺伝子の前記ゲノム配列情報の位置と未知の遺伝子のゲノム配列上の位置との距離を算出する算出工程と、

前記算出工程によって算出された距離の近い順に前記周辺遺伝子をソートするソート工程と、

前記発現部位情報取得工程によって取得された発現部位に関する情報を、前記ソート工程によってソートされた周辺遺伝子の順に出力する発現部位情報出力工程と、

を含んだことを特徴とする遺伝子発現部位予測支援方法。

【請求項 3】 前記発現部位情報出力工程は、ソートされた上位の周辺遺伝子によって特定されている発現部位に関する情報がある場合は、当該発現部位に

関する情報以外の情報のみを出力することを特徴とする請求項 2 に記載の遺伝子発現部位予測支援方法。

【請求項 4】 前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭の位置と、前記周辺遺伝子のゲノム配列上の先頭の位置との距離を算出することを特徴とする請求項 1 ～ 3 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【請求項 5】 前記算出工程は、前記未知の遺伝子のゲノム配列上の最後尾の位置と、前記周辺遺伝子のゲノム配列上の最後尾の位置との距離を算出することを特徴とする請求項 1 ～ 3 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【請求項 6】 前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭の位置と、前記周辺遺伝子のゲノム配列上の最後尾の位置との距離を算出することを特徴とする請求項 1 ～ 3 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【請求項 7】 前記算出工程は、前記未知の遺伝子のゲノム配列上の最後尾の位置と、前記周辺遺伝子のゲノム配列上の先頭の位置との距離を算出することを特徴とする請求項 1 ～ 3 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【請求項 8】 前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置と、前記周辺遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置との距離を算出することを特徴とする請求項 1 ～ 3 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【請求項 9】 ゲノム配列上における、発現部位が特定されている周辺遺伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出させる算出工程と、

前記算出工程によって算出された距離に基づいて、前記未知の遺伝子の発現部位を特定させる特定工程と、

をコンピュータに実行させることを特徴とする遺伝子発現部位予測支援プログラム。

【請求項 1 0】 ゲノム配列上における、発現部位が特定されている周辺遺

伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出する算出手段と、

前記算出手段によって算出された距離に基づいて、前記未知の遺伝子の発現部位を特定する特定手段と、

を備えたことを特徴とする遺伝子発現部位予測支援装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、未知の遺伝子配列が与えられたときに、当該遺伝子が発現する部位の予測を支援する遺伝子発現部位予測支援方法、遺伝子発現部位予測支援プログラム、および遺伝子発現部位予測支援装置に関する。

【0002】

【従来の技術】

近年の遺伝子工学の進歩に伴い遺伝子配列の解読が進んでおり、それに伴って増加している未知遺伝子の機能解明が求められている。遺伝子の機能解明に有効な方法として、対象遺伝子が発現している部位を基にした機能予測がある。そのため、未知遺伝子の発現部位を予測する技術が必要とされる。

【0003】

従来、未知遺伝子の発現部位を予測するには、実際に実験により未知遺伝子が発現している組織を探索する方法と、計算機により大量のEST配列に対して未知遺伝子配列のホモロジー検索をおこなう方法とが存在する。これは、ESTデータベースから、発現部位を予測したい遺伝子配列のホモロジーを検索し、当該ESTの発現情報を抽出するものである。

【0004】

【発明が解決しようとする課題】

しかしながら、上記の従来技術のうち、実際に実験により未知遺伝子が発現している組織を探索する方法においては非常な労力を要し、効率的な機能解析をおこなうことができないという問題点があった。また、ホモロジー検索をおこなう方法においても、大量のEST配列の管理あるいはホモロジー検索そのものの処

理に時間がかかるため、やはり同様に迅速な機能解析を妨げているという問題点があった。

【0005】

この発明は上記問題を解決するため、未知遺伝子の発現部位予測・機能解析を迅速かつ効率的におこなうことが可能な遺伝子発現部位予測支援方法、遺伝子発現部位予測支援プログラム、遺伝子発現部位予測支援装置を提供することを目的とする。

【0006】

【課題を解決するための手段】

上述した課題を解決し、目的を達成するため、この発明にかかる遺伝子発現部位予測支援方法、遺伝子発現部位予測支援プログラムおよび遺伝子発現部位予測支援装置は、ゲノム配列上における、発現部位が特定されている周辺遺伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出し、算出された距離に基づいて、前記未知の遺伝子の発現部位を特定することを特徴とする。

【0007】

これらの発明によれば、ゲノム配列上で遺伝子間の距離が近いもの程同一部位に発現する傾向を用いて、周辺遺伝子の発現部位情報から未知の遺伝子の発現部位を推測することができる。

【0008】

【発明の実施の形態】

以下に添付図面を参照して、この発明にかかる遺伝子発現部位予測支援方法、遺伝子発現部位予測支援プログラムおよび遺伝子発現部位予測支援装置の好適な実施の形態を詳細に説明する。

【0009】

(遺伝子発現部位予測支援方法の概要)

まず、この発明の本実施の形態にかかる遺伝子発現部位予測支援方法の概要について説明する。図1は、この発明の本実施の形態にかかる遺伝子発現部位予測支援方法の概要を示す説明図である。

【0010】

図1において、100はゲノム配列であり、101は、未知遺伝子（対象遺伝子）であり、102および103は周辺遺伝子である。未知遺伝子101をゲノム配列100上にmappingし、周辺遺伝子102、103の発現部位を抽出し、未知遺伝子101との距離（距離a、距離b）によってスコア付けをおこなって、未知遺伝子101の発現部位の予測に用いる。

【0011】

周辺遺伝子A102にあつては、発現部位はbrain（脳）およびovary（卵巣）であり、未知遺伝子101との距離はaである。一方、周辺遺伝子B103にあつては、発現部位はspleen（脾臓）であり、未知遺伝子101との距離はbである。距離a<距離bであるから、この未知遺伝子101の発現部位としては、brain（脳）またはovary（卵巣）に、より関連があると予測することができる。

【0012】

図2は、ゲノム配列上の遺伝子と発現部位の傾向を示す図表であり、同一部位に発現している遺伝子の組数（ヒト21番染色体）についてグラフ化したものである。この図表からもわかるように、ゲノム配列100上で遺伝子間の距離が近いものほど同一部位に発現する傾向があるといえる。この発明の本実施の形態にかかる遺伝子発現部位予測支援方法では、この傾向に基づいて、未知遺伝子101の発現部位の予測をおこなうものである。

【0013】

（遺伝子発現部位予測支援装置のハードウェア構成）

つぎに、この発明の本実施の形態にかかる遺伝子発現部位予測支援装置のハードウェア構成について説明する。図3は、この発明の本実施の形態にかかる遺伝子発現部位予測支援装置のハードウェア構成の一例を示すブロック図である。

【0014】

図3において、遺伝子発現部位予測支援装置は、CPU301と、ROM302と、RAM303と、HDD304と、HD305と、FDD（フレキシブルディスクドライブ）306と、着脱可能な記録媒体の一例としてのFD（フレキシブルディスク）307と、ディスプレイ308と、I/F（インタフェース）

309と、キーボード311と、マウス312と、スキャナ313と、プリンタ314と、を備えている。また、各構成部はバス300によってそれぞれ接続されている。

【0015】

ここで、CPU301は、遺伝子発現部位予測支援装置の全体の制御を司る。ROM302は、ブートプログラムなどのプログラムを記憶している。RAM303は、CPU301のワークエリアとして使用される。HDD304は、CPU301の制御にしたがってHD305に対するデータのリード／ライトを制御する。HD305は、HDD304の制御で書き込まれたデータを記憶する。

【0016】

FDD306は、CPU301の制御にしたがってFD307に対するデータのリード／ライトを制御する。FD307は、FDD306の制御で書き込まれたデータを記憶したり、FD307に記録されたデータを情報処理装置へ読み取らせたりする。着脱可能な記録媒体として、FD307のほか、CD-ROM (CD-R、CD-RW)、MO、DVD (Digital Versatile Disk)、メモリーカードなどであってもよい。ディスプレイ308は、カーソル、アイコンあるいはツールボックスをはじめ、文書、画像、機能情報などのデータを表示する。たとえば、CRT、TFT液晶ディスプレイ、プラズマディスプレイなどである。

【0017】

I/F (インタフェース) 309は、通信回線を通じてLANやインターネットなどのネットワーク310に接続され、ネットワーク310を介して、データベースなどを備えた他のサーバーや情報処理装置に接続される。そして、I/F 309は、ネットワーク310と内部とのインタフェースを司り、他のサーバーや情報端末装置からのデータの入出力を制御する。I/F 309は、たとえばモデムなどである。

【0018】

キーボード311は、文字、数字、各種指示などの入力のためのキーを備え、データの入力をおこなう。タッチパネル式の入力パッドやテンキーなどであって

もよい。マウス 312 は、カーソルの移動や範囲選択、あるいはウインドウの移動やサイズの変更などをおこなう。ポインティングデバイスとして同様の機能を備えるものであれば、トラックボール、ジョイスティックなどであってもよい。

【0019】

スキャナ 313 は、図形や写真などの画像を光学的に読み取り、情報処理装置内に画像データとして取り込む。さらに OCR 機能も備えており、OCR 機能によって、印刷された情報を読み取ってデータ化することもできる。また、プリンタ 314 は、輪郭画像情報などの画像データや文書データを印刷する。たとえば、レーザプリンタ、インクジェットプリンタなどである。

【0020】

(遺伝子発現部位予測支援装置の機能的構成)

つぎに、遺伝子発現部位予測支援装置の機能的構成について説明する。図 4 は、この発明の本実施の形態にかかる遺伝子発現部位予測支援装置の機能的構成の一例を示すブロック図である。図 4 において、401 は未知遺伝子配列読込部であり、402 はゲノム配列読込部であり、403 は配列比較処理部であり、404 は未知遺伝子配列位置取得部であり、405 は周辺遺伝子探索処理部であり、406 は発現プロファイル DB 読込部であり、407 は周辺遺伝子発現部位取得部であり、408 は周辺遺伝子発現部位重み付け処理部であり、409 は出力部である。

【0021】

ここで、未知遺伝子配列読込部 401 は、未知遺伝子 101 の配列情報を読み込むことによって未知遺伝子 101 の配列情報の入力を受け付ける。未知遺伝子配列読込部 401 は、I/F 309 あるいは、キーボード 311、マウス 312、スキャナ 313 などによってその機能を実現することができる。

【0022】

また、ゲノム配列読込部 402 は、未知遺伝子 101 に関連あるゲノム配列 100 の情報を読み込むことによって未知遺伝子 101 の配列情報の入力を受け付ける。ゲノム配列読込部 402 も、未知遺伝子配列読込部 401 と同様に、I/F 309 あるいは、キーボード 311、マウス 312、スキャナ 313 などによ

ってその機能を実現することができる。

【0023】

また、配列比較処理部403は、未知遺伝子配列読込部401において読み込まれた未知遺伝子101の配列情報と、ゲノム配列読込部402において読み込まれたゲノム配列100の情報とを比較することによって、未知遺伝子配列読込部401において読み込まれた未知遺伝子101の配列情報を、ゲノム配列読込部402において読み込まれたゲノム配列100の情報にmappingする。mappingについての詳細は後述する。

【0024】

また、未知遺伝子配列位置取得部404は、配列比較処理部403によって比較された比較結果451、すなわちゲノム配列100上にmappingされた未知遺伝子101のゲノム配列100上の位置を取得する。そして取得された位置については未知遺伝子位置情報452として周辺遺伝子探索処理部405へ送られる。

【0025】

また、周辺遺伝子探索処理部405は、未知遺伝子101のゲノム配列100上の位置の周辺に存在する、発現部位が特定されている周辺遺伝子（遺伝子A102，遺伝子B103）を、ゲノム配列読込部402から送られた周辺遺伝子名・位置情報453に基づいて検索する。そして、その結果として、周辺遺伝子名情報454を発現プロファイルDB読込部406へ送るとともに、周辺遺伝子位置情報456を、周辺遺伝子発現部位重み付け処理部408へ送る。

【0026】

また、発現プロファイルDB読込部406は、周辺遺伝子探索処理部405によって検索され、その周辺遺伝子名情報454が送られた周辺遺伝子についての発現プロファイルを、後述する発現プロファイルDB553から読み込む。

【0027】

また、周辺遺伝子発現部位取得部407は、発現プロファイルDB読込部406によって読み込まれた発現プロファイルの中から、当該周辺遺伝子によって特定されている発現部位に関する情報を取得し、周辺遺伝子発現部位情報455と

して、周辺遺伝子発現部位重み付け処理部 408 へ送る。周辺遺伝子発現部位取得部 407 は、I/F 309 あるいは、キーボード 311、マウス 312、スキャナ 313 などによってその機能を実現することができる。

【0028】

また、周辺遺伝子発現部位重み付け処理部 408 は、ゲノム配列 100 上における、発現部位が特定されている周辺遺伝子 102、103 が存在する位置と未知遺伝子 101 が存在する位置との距離を算出し、算出された距離に基づいて、未知遺伝子 101 の発現部位を特定する。たとえば、算出された距離が短い一つまたは複数の周辺遺伝子の発現部位を未知遺伝子 101 の発現部位として特定することができる。

【0029】

また、発現部位が特定されている周辺遺伝子のゲノム配列情報の位置と未知遺伝子のゲノム配列上の位置との距離を算出し、算出された距離の近い順に周辺遺伝子をソートする。そして、ソートされた上位の周辺遺伝子によって特定されている発現部位に関する情報がある場合は、その情報をマージする。すなわち、当該発現部位に関する情報を削除し、当該発現部位に関する情報以外の情報のみを残し、後述する出力部 409 が、残された当該発現部位に関する情報以外の情報のみ出力する。

【0030】

なお、配列比較処理部 403、未知遺伝子配列位置取得部 404、周辺遺伝子探索処理部 405 は、周辺遺伝子発現部位取得部 407、周辺遺伝子発現部位重み付け処理部 408 は、たとえば、図 3 に示す ROM 302、RAM 303、HD 305 あるいは FD 307 に記憶されたプログラムを CPU 301 が実行することによってそれらの機能を実現する。

【0031】

また、出力部 409 は、周辺遺伝子発現部位重み付け処理部 408 によって処理された発現部位に関する情報を、ソートされた周辺遺伝子の順に出力する。また、取得された発現部位に関する情報を一覧表示するようにしてもよい。これが未知遺伝子予測発現部位一覧 457 となる。出力部 409 は、たとえば図 3 に示

したFD307およびFDD306、I/F309などによって外部へ出力することができる。また、プリンタ314などによって印刷させることもでき、さらには、ディスプレイ308などによって表示させるようにすることもできる。

【0032】

(遺伝子発現部位予測支援装置の処理手順)

つぎに、遺伝子発現部位予測支援装置の処理の手順について説明する。図5は、この発明の本実施の形態にかかる遺伝子発現部位予測支援装置の処理の手順を示すフローチャートである。

【0033】

図5のフローチャートにおいて、まず、(1)対象遺伝子(未知遺伝子)の配列を配列Aのデータベース551から読み取ることによって入力する(ステップS501)。つぎに(2)ゲノム配列情報を配列Bのデータベース552から読み取ることによって対象遺伝子のゲノム配列へのmappingをおこなう(ステップS502)。そして、(3)対象遺伝子と周辺に存在する周辺遺伝子の距離を算出する(ステップS503)。さらに、(4)対象遺伝子の周辺に存在する周辺遺伝子の発現部位を発現プロファイルDB553から抽出する(ステップS504)。そして、(5)周辺遺伝子の発現部位を対象遺伝子との距離で重み付けして出力する(ステップS505)。

【0034】

(1)対象遺伝子(未知遺伝子)の配列入力(ステップS501)：

図6は、未知遺伝子101の配列A551の配列情報の一例を示す説明図である。この配列A551の配列情報を入力する。入力方法としては、ネットワーク310を介して入力してもよく、記録メディアを介して入力してもよい。また、キーボード311を用いて直接入力する方法であってもよく、さらにはOCR機能を搭載したスキャナ313を用いてイメージ情報としての配列情報を入力し、デジタル情報に変換することによって入力する方法であってもよい。

【0035】

(2)未知遺伝子のゲノム配列へのmapping(ステップS502)：

図7は、ゲノム配列(配列B552)の配列情報の一部を示す説明図である。

未知遺伝子 101 の配列 A 551 と、図 7 に示すゲノム配列 B 552 の類似領域をホモロジー検索により求めた結果を図 8 に示す。このとき、配列 A の配列 B との関係は、以下の通りであることが求められる。

【0036】

- ① 配列 A 551 の殆どが配列 B 552 上に mapping されていること。
- ② 配列 A 551 が配列 B 552 に mapping されている状態で、配列 A 551 には隙間 (gap) 801 があってもよい。
- ③ 配列 A 551 に gap がある場合、gap によって区切られた配列 A 551 の断片の配列 B 552 上での mapping 順序は、配列 A 551 上での当該断片の出現順序と同じであること。

【0037】

図 8 からわかるように、対象遺伝子の配列 A 551 のゲノム配列 B 552 上の位置 (先頭位置) は、「12313789 (bp)」802 となる。

【0038】

(3) 未知遺伝子と周辺遺伝子の距離の算出 (ステップ S 503) :

図 9 は、ゲノム配列 B 上に mapping された未知遺伝子の配列 A の周辺遺伝子とその位置情報の一例を示す説明図である。図 9 において、左欄には周辺遺伝子名が記載されるとともに、右欄には当該周辺遺伝子の位置 (先頭位置) が記載されている。さらに、周辺遺伝子はゲノム配列 100 上の出現順にソートされて記載されている。なお、ソートはステップ S 505 の出力の際おこなうようにしてもよい。

【0039】

また、図 10 は、以下の計算式により、未知遺伝子の配列 A と周辺遺伝子の距離を算出した結果を示す説明図である。

【0040】

未知遺伝子の配列 A と周辺遺伝子の距離 (bp) =
| (対象配列 A のゲノム配列 B 上の位置 (bp))
- (周辺遺伝子のゲノム配列 B 上の位置 (bp)) |

【0041】

図10において、左欄には周辺遺伝子名が記載されるとともに、右欄には当該周辺遺伝子の未知遺伝子からの距離が記載されている。さらに、周辺遺伝子は前記距離の短い順にソートされて記載されている。具体的には、周辺遺伝子「C21orf42」は、図9によると、ゲノム配列B552上の位置が「12337804 (bp)」であり、対象遺伝子の配列A551のゲノム配列B552上の位置は、「12313789 (bp)」であるから、未知遺伝子からの距離は、 $|12313789 (bp) - 12337804 (bp)| = 24015 (bp)$ となる。また、周辺遺伝子「ADAMTS1」は、図9によると、ゲノム配列B上の位置が「13788256 (bp)」であるから、未知遺伝子からの距離は、 $|12313789 (bp) - 13788256 (bp)| = 1474467 (bp)$ となる。

【0042】

上述のように、発現部位が特定されている周辺遺伝子102, 103が存在する位置と未知遺伝子101が存在する位置との距離を算出する際、両遺伝子のゲノム配列100上の先頭位置に基づいておこなったが、これに限定されるものではない。たとえば、第2の算出方法として、未知遺伝子101のゲノム配列100上の最後尾の位置と、周辺遺伝子のゲノム配列100上の最後尾の位置との距離を算出するようにしてもよい。

【0043】

さらに、未知遺伝子101のゲノム配列100上の先頭の位置と、周辺遺伝子のゲノム配列上の最後尾の位置との距離を算出するようにしてもよい。この場合、周辺遺伝子が未知遺伝子101よりもゲノム配列100上前方に存在しても、後方に存在してもいずれでもよい。また、その逆に、未知遺伝子101のゲノム配列100上の最後尾の位置と、周辺遺伝子のゲノム配列100上の先頭の位置との距離を算出するようにしてもよい。この場合も、周辺遺伝子が未知遺伝子101よりもゲノム配列100上前方に存在しても、後方に存在してもいずれでもよい。

【0044】

また、遺伝子の位置を特定するのに、遺伝子の先頭または最後尾を基準とする

代わりに、遺伝子のゲノム配列 100 上の先頭と最後尾との中間の任意の位置を基準として遺伝子間の距離を算出するようにしてもよい。より具体的には、遺伝子の先頭と最後尾の中間に該当する部分を位置の基準とすることができる。たとえば、未知遺伝子 101 のゲノム配列 100 上の先頭と最後尾との中間の任意の位置と、周辺遺伝子のゲノム配列 100 上の先頭と最後尾との中間の任意の位置との距離を算出するとよい。

【0045】

さらに、中間の任意の位置と、先頭または最後尾との距離を算出することもできる。具体的には、未知遺伝子 101 のゲノム配列 100 上の先頭と最後尾との中間の任意の位置と、前記周辺遺伝子のゲノム配列 100 上の先頭の位置との距離を算出する。また、未知遺伝子 101 のゲノム配列 100 上の先頭と最後尾との中間の任意の位置と、前記周辺遺伝子のゲノム配列 100 上の最後尾の位置との距離を算出することもできる。

【0046】

同様に、未知遺伝子 101 のゲノム配列 100 上の先頭の位置と、周辺遺伝子のゲノム配列 100 上の先頭と最後尾との中間の任意の位置との距離を算出することもでき、未知遺伝子 101 のゲノム配列 100 上の最後尾の位置と、前記周辺遺伝子のゲノム配列 100 上の先頭と最後尾との中間の任意の位置との距離を算出することもできる。

【0047】

(4) 未知遺伝子の周辺に存在する周辺遺伝子の発現部位の抽出（ステップ S504）：

図 11 は、発現プロファイル DB の例として、周辺遺伝子「ADAMTS1」の内容を示す説明図である。図 11 において、GENE1101 は遺伝子略称を示しており、また、EXPRESS1102 は発現部位の情報を示している。その他の各項目については、本実施の形態とは直接的な関連性を持たないためそれらの説明は省略する。

【0048】

また、図 12 は、図 6 に示した各遺伝子の発現部位を発現プロファイル DB 5

53 から抽出した結果（リスト）を示す説明図である。図12において、左欄には周辺遺伝子名が記載されるとともに、右欄には発現部位の内容が記載されている。さらに、周辺遺伝子は、図10と同様に、前記距離の短い順にソートされて記載されている。なお、ソートはステップS505の出力の際おこなうようにしてもよい。

【0049】

(5) 周辺遺伝子の発現部位を未知遺伝子との距離で重み付け・出力（ステップS505）：

図12に示すリストで、上位に発現している部位と同一の発現部位を持つ下位の発現部位をリスト上から削除する。たとえば、図12より、未知遺伝子の最近傍の周辺遺伝子「C21orf42」の下位に位置する周辺遺伝子「MRPL39」の発現部位として、「testis」が挙げられているが、これは、上位の周辺遺伝子「C21orf42」で既に発現していることがわかっているため、下位の周辺遺伝子「MRPL39」の項では削除する。本操作を最下位の周辺遺伝子の発現部位まで繰り返す。削除した結果を図13に示す。

【0050】

上記(1)～(5)の処理をおこなうことによって、未知遺伝子の配列Aについて図13に示すリストが得られる。本リストから、未知遺伝子の配列Aが発現していると予測される部位を容易に推測することができる。

【0051】

以上説明したように、本実施の形態によれば、未知遺伝子配列から発現部位を容易に予測することができ、実験コストおよび計算機資源を節約することができる。したがって、本発明を利用することで、急増している未知遺伝子の機能解明を迅速におこなうことが可能となり、遺伝子工学の発展に寄与することが大きい。

【0052】

なお、本実施の形態における遺伝子発現部位予測支援方法は、あらかじめ用意されたコンピュータ読み取り可能なプログラムであってもよく、またそのプログラムをパーソナルコンピュータやワークステーションなどのコンピュータで実行

することによって実現される。このプログラムは、HD、FD、CD-ROM、MO、DVDなどのコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。また、このプログラムは、インターネットなどのネットワークを介して配布することが可能な伝送媒体であってもよい。

【0053】

(付記1) ゲノム配列上における、発現部位が特定されている周辺遺伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出する算出工程と、

前記算出工程によって算出された距離に基づいて、前記未知の遺伝子の発現部位を特定する特定工程と、

を含んだことを特徴とする遺伝子発現部位予測支援方法。

【0054】

(付記2) 前記特定工程は、前記算出工程によって算出された距離が短い一つまたは複数の周辺遺伝子の発現部位を前記未知の遺伝子の発現部位として特定することを特徴とする付記1に記載の遺伝子発現部位予測支援方法。

【0055】

(付記3) 未知の遺伝子の配列情報を入力する入力工程と、

前記入力工程によって入力された配列情報にかかる未知の遺伝子のゲノム配列上の位置を取得する位置情報取得工程と、

前記位置情報取得工程によって取得された未知の遺伝子のゲノム配列上の位置の周辺に存在する、発現部位が特定されている周辺遺伝子を検索する検索工程と、

前記検索工程によって検索された周辺遺伝子によって特定されている発現部位に関する情報を取得する発現部位情報取得工程と、

前記検索工程によって検索された発現部位が特定されている周辺遺伝子の前記ゲノム配列情報の位置と未知の遺伝子のゲノム配列上の位置との距離を算出する算出工程と、

前記算出工程によって算出された距離の近い順に前記周辺遺伝子をソートするソート工程と、

前記発現部位情報取得工程によって取得された発現部位に関する情報を、前記ソート工程によってソートされた周辺遺伝子の順に出力する発現部位情報出力工程と、

を含んだことを特徴とする遺伝子発現部位予測支援方法。

【0056】

（付記4）前記発現部位情報出力工程は、ソートされた上位の周辺遺伝子によって特定されている発現部位に関する情報がある場合は、当該発現部位に関する情報以外の情報のみを出力することを特徴とする付記3に記載の遺伝子発現部位予測支援方法。

【0057】

（付記5）前記発現部位情報出力工程は、前記発現部位情報取得工程によって取得された発現部位に関する情報を一覧表示することを特徴とする付記3または4に記載の遺伝子発現部位予測支援方法。

【0058】

（付記6）前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭の位置と、前記周辺遺伝子のゲノム配列上の先頭の位置との距離を算出することを特徴とする付記1～5のいずれか一つに記載の遺伝子発現部位予測支援方法。

【0059】

（付記7）前記算出工程は、前記未知の遺伝子のゲノム配列上の最後尾の位置と、前記周辺遺伝子のゲノム配列上の最後尾の位置との距離を算出することを特徴とする付記1～5のいずれか一つに記載の遺伝子発現部位予測支援方法。

【0060】

（付記8）前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭の位置と、前記周辺遺伝子のゲノム配列上の最後尾の位置との距離を算出することを特徴とする付記1～5のいずれか一つに記載の遺伝子発現部位予測支援方法。

【0061】

（付記9）前記算出工程は、前記未知の遺伝子のゲノム配列上の最後尾の位置と、前記周辺遺伝子のゲノム配列上の先頭の位置との距離を算出することを特徴とする付記1～5のいずれか一つに記載の遺伝子発現部位予測支援方法。

【 0 0 6 2 】

（付記 1 0）前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置と、前記周辺遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置との距離を算出することを特徴とする付記 1 ～ 5 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【 0 0 6 3 】

（付記 1 1）前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置と、前記周辺遺伝子のゲノム配列上の先頭の位置との距離を算出することを特徴とする付記 1 ～ 5 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【 0 0 6 4 】

（付記 1 2）前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置と、前記周辺遺伝子のゲノム配列上の最後尾の位置との距離を算出することを特徴とする付記 1 ～ 5 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【 0 0 6 5 】

（付記 1 3）前記算出工程は、前記未知の遺伝子のゲノム配列上の先頭の位置と、前記周辺遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置との距離を算出することを特徴とする付記 1 ～ 5 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【 0 0 6 6 】

（付記 1 4）前記算出工程は、前記未知の遺伝子のゲノム配列上の最後尾の位置と、前記周辺遺伝子のゲノム配列上の先頭と最後尾との中間の任意の位置との距離を算出することを特徴とする付記 1 ～ 5 のいずれか一つに記載の遺伝子発現部位予測支援方法。

【 0 0 6 7 】

（付記 1 5）ゲノム配列上における、発現部位が特定されている周辺遺伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出させる算出工程と、

前記算出工程によって算出された距離に基づいて、前記未知の遺伝子の発現部

位を特定させる特定工程と、

をコンピュータに実行させることを特徴とする遺伝子発現部位予測支援プログラム。

【0068】

(付記16) 前記特定工程は、前記算出工程によって算出された距離が短い一つまたは複数の周辺遺伝子の発現部位を前記未知の遺伝子の発現部位として特定させることを特徴とする付記15に記載の遺伝子発現部位予測支援プログラム。

【0069】

(付記17) 未知の遺伝子の配列情報の入力させる入力工程と、

前記入力工程によって入力された配列情報にかかる未知の遺伝子のゲノム配列上の位置を取得させる位置情報取得工程と、

前記位置情報取得工程によって取得された未知の遺伝子のゲノム配列上の位置の周辺に存在する、発現部位が特定されている周辺遺伝子を検索させる検索工程と、

前記検索工程によって検索された周辺遺伝子によって特定されている発現部位に関する情報を取得させる発現部位情報取得工程と、

前記検索工程によって検索された発現部位が特定されている周辺遺伝子の前記ゲノム配列情報の位置と未知の遺伝子のゲノム配列上の位置との距離を算出させる算出工程と、

前記算出工程によって算出された距離の近い順に前記周辺遺伝子をソートさせるソート工程と、

前記発現部位情報取得工程によって取得された発現部位に関する情報を、前記ソート工程によってソートされた周辺遺伝子の順に出力させる発現部位情報出力工程と、

をコンピュータに実行させることを特徴とする遺伝子発現部位予測支援プログラム。

【0070】

(付記18) ゲノム配列上における、発現部位が特定されている周辺遺伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出する算出手段と、

前記算出手段によって算出された距離に基づいて、前記未知の遺伝子の発現部位を特定する特定手段と、

を備えたことを特徴とする遺伝子発現部位予測支援装置。

【0071】

(付記19) 前記特定手段は、前記算出手段によって算出された距離が短い一つまたは複数の周辺遺伝子の発現部位を前記未知の遺伝子の発現部位として特定することを特徴とする付記18に記載の遺伝子発現部位予測支援装置。

【0072】

(付記20) 未知の遺伝子の配列情報の入力を受け付ける入力手段と、

前記入力手段によって入力された配列情報にかかる未知の遺伝子のゲノム配列上の位置を取得する位置情報取得手段と、

前記位置情報取得手段によって取得された未知の遺伝子のゲノム配列上の位置の周辺に存在する、発現部位が特定されている周辺遺伝子を検索する検索手段と、

前記検索手段によって検索された周辺遺伝子によって特定されている発現部位に関する情報を取得する発現部位情報取得手段と、

前記検索手段によって検索された発現部位が特定されている周辺遺伝子の前記ゲノム配列情報の位置と未知の遺伝子のゲノム配列上の位置との距離を算出する算出手段と、

前記算出手段によって算出された距離の近い順に前記周辺遺伝子をソートするソート手段と、

前記発現部位情報取得手段によって取得された発現部位に関する情報を、前記ソート手段によってソートされた周辺遺伝子の順に出力する発現部位情報出力手段と、

を備えたことを特徴とする遺伝子発現部位予測支援装置。

【0073】

【発明の効果】

以上説明したように、この発明によれば、ゲノム配列上で遺伝子間の距離が近いもの程同一部位に発現する傾向を用いて、周辺遺伝子の発現部位情報から未知

の遺伝子の発現部位を推測することができ、これによって、未知遺伝子の発現部位予測・機能解析を迅速かつ効率的におこなうことが可能な遺伝子発現部位予測支援方法、遺伝子発現部位予測支援プログラム、遺伝子発現部位予測支援装置が得られるという効果を奏する。

【図面の簡単な説明】

【図 1】

この発明の本実施の形態にかかる遺伝子発現部位予測支援方法の概要を示す説明図である。

【図 2】

ゲノム配列上の遺伝子と発現部位の傾向を示す図表である。

【図 3】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置のハードウェア構成の一例を示すブロック図である。

【図 4】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置の機能的構成の一例を示すブロック図である。

【図 5】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置の処理の手順を示すフローチャートである。

【図 6】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置において用いられる未知遺伝子の配列 A の配列情報の一例を示す説明図である。

【図 7】

ゲノム配列（配列 B）の配列情報の一部を示す説明図である。

【図 8】

配列 A と配列 B との比較結果を示す説明図である。

【図 9】

周辺遺伝子とその位置情報の一例を示す説明図である。

【図 10】

未知遺伝子の配列 A と周辺遺伝子の距離を算出した結果を示す説明図である。

【図 1 1】

発現プロファイルデータベースの内容の一例を示す説明図である。

【図 1 2】

周辺遺伝子の発現部位情報の一例を示す説明図である。

【図 1 3】

周辺遺伝子の発現部位情報の別の一例を示す説明図である。

【符号の説明】

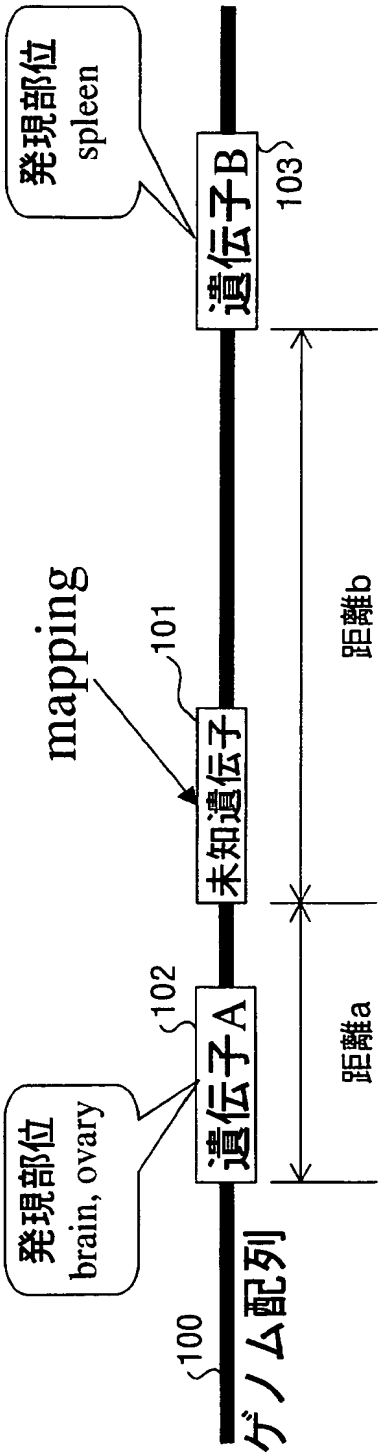
- 1 0 0 ゲノム配列
- 1 0 1 未知遺伝子（対象遺伝子）
- 1 0 2, 1 0 3 周辺遺伝子
- 4 0 1 未知遺伝子配列読込部
- 4 0 2 ゲノム配列読込部
- 4 0 3 配列比較処理部
- 4 0 4 未知遺伝子配列位置取得部
- 4 0 5 周辺遺伝子探索処理部
- 4 0 6 発現プロファイル DB 読込部
- 4 0 7 周辺遺伝子発現部位取得部
- 4 0 8 周辺遺伝子発現部位重み付け処理部
- 4 0 9 出力部
- 5 5 1 配列 A
- 5 5 2 配列 B
- 5 5 3 発現プロファイル DB

【書類名】

図面

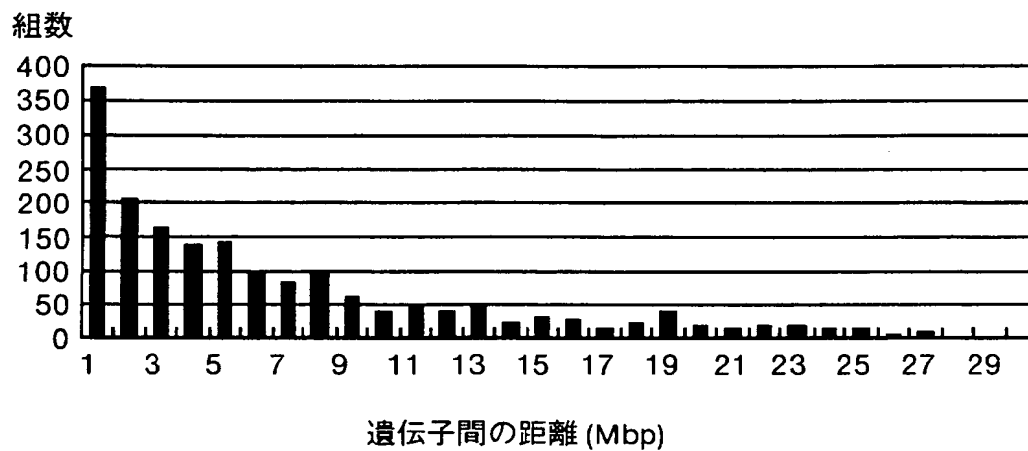
【図 1】

この発明の本実施の形態にかかる遺伝子発現部位予測支援方法の概要を示す説明図



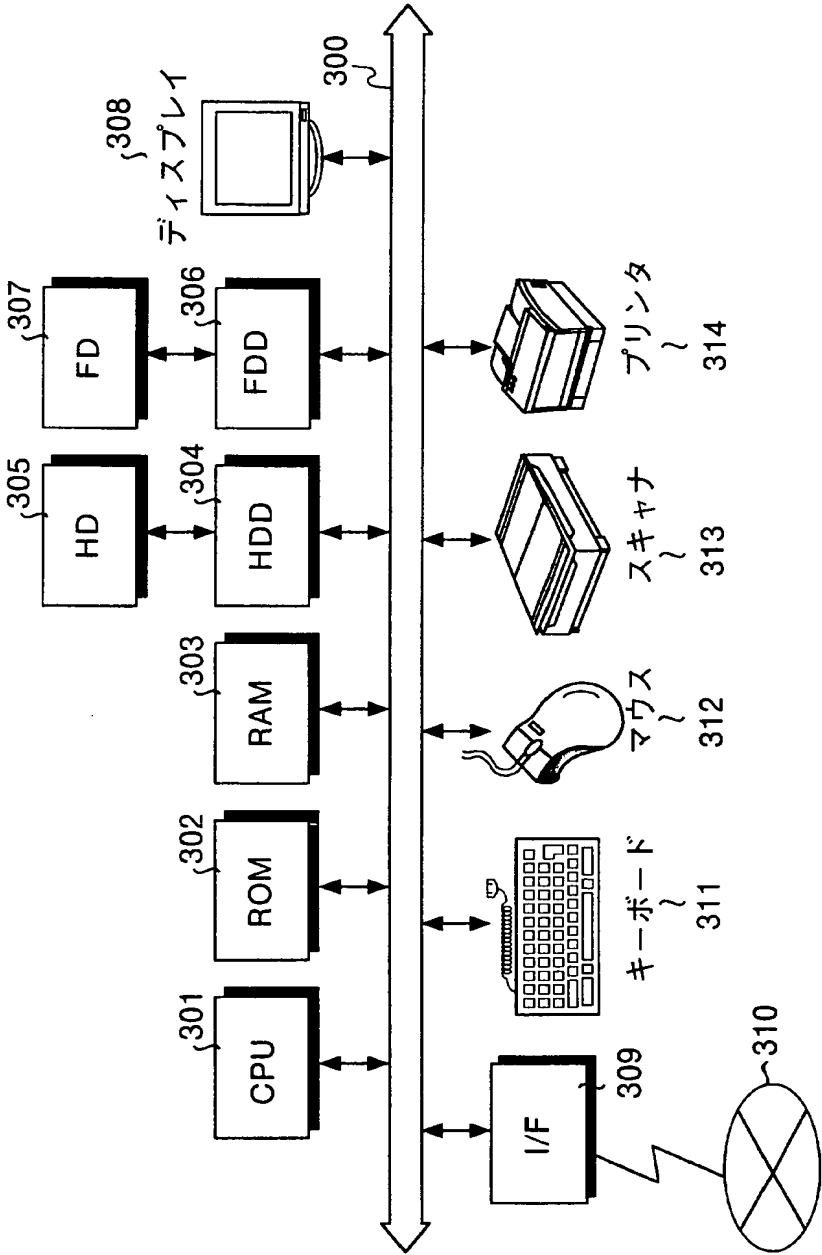
【図 2】

ゲノム配列上の遺伝子と発現部位の傾向を示す図表



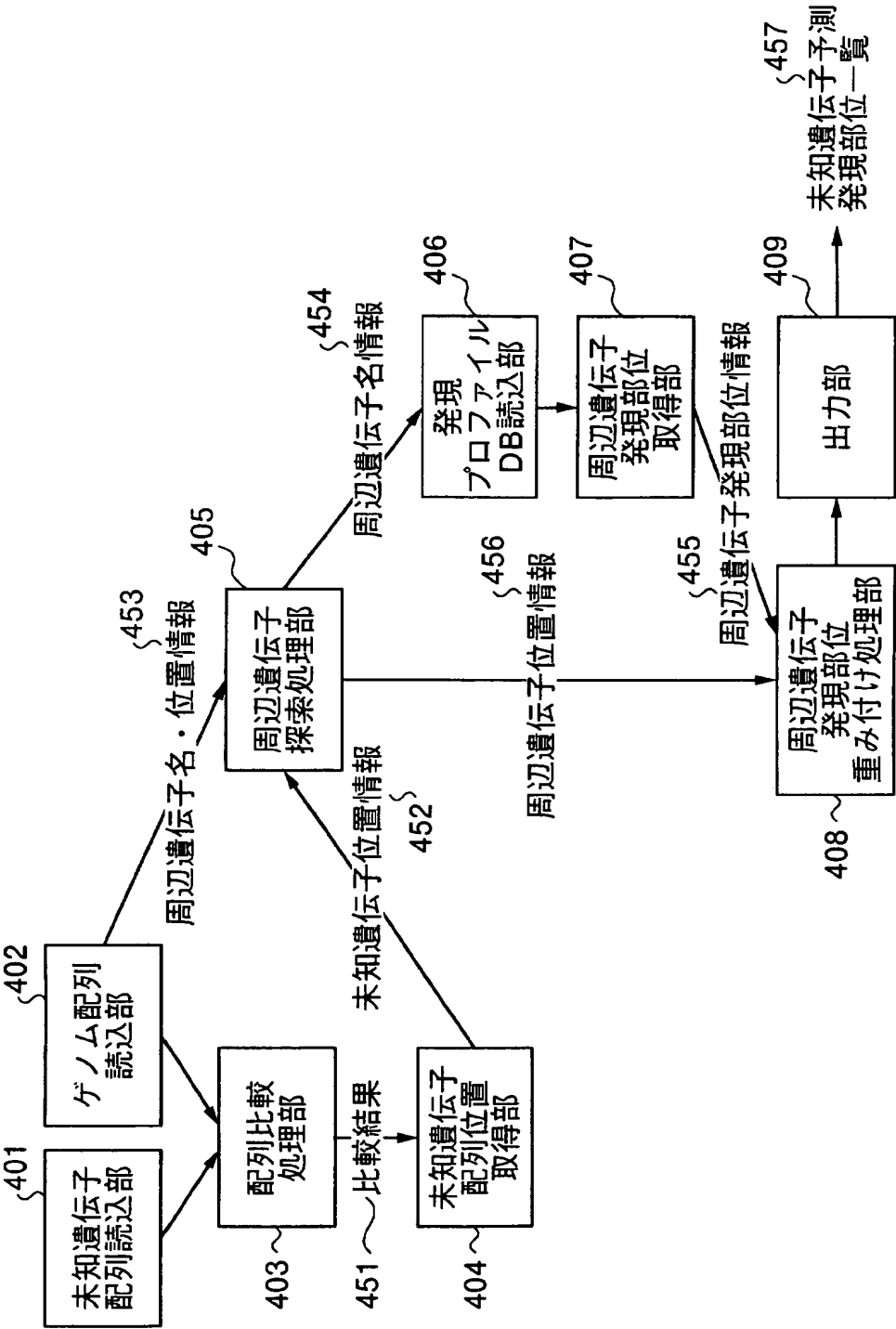
【図 3】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置のハードウェア構成の一例を示すブロック図



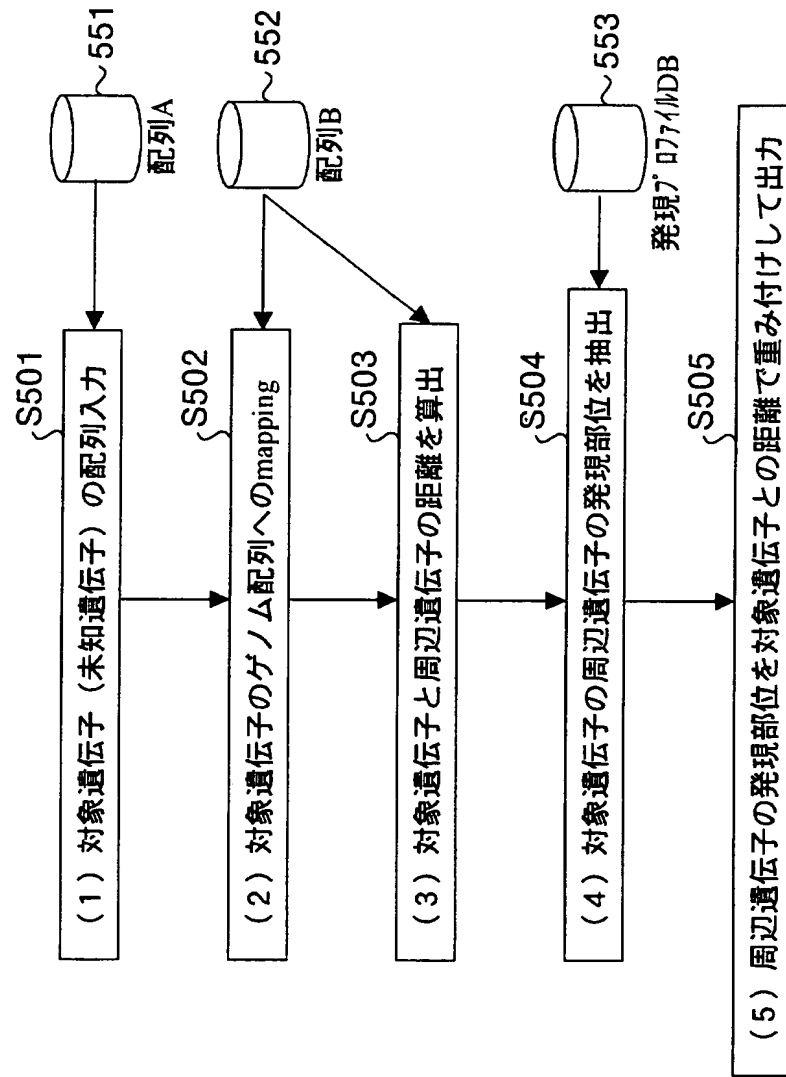
【図 4】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置の機能的構成の一例を示すブロック図



【図 5】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置の
処理の手順を示すフローチャート



【図 6】

この発明の本実施の形態にかかる遺伝子発現部位予測支援装置において
用いられる未知遺伝子の配列 A の配列情報の一例を示す説明図

551

```
GGCTGCCGAAGATGGCGGAGGTGCAGGTCCTGGTGCATGGTCGATGGTCGAGGCCATCTCCTGGTCCGCCTGGC
GGCCATCGTGGCTAAACAGGTACTGCTGGCCGGAAAGTGGTGGTCGTACGCTGCGAAGGCATCAACATT
TCTGGCAATTTCTACAGAAACAAGTTGAAGTACCTGGGTTTCTCCGCAAGCGGATGAACACCCACCTTT
CCCGAGTCCCTACCATTCCGGGCCCCCAGCCGCATCTTCTGGCGGACCGTGCGAGGTATGCCGCCCC
ACAAGACCAAGCGAGGCCAGGCTTCTTGACCCGCTCAAGGTGTTTGACCCGCATCCACCCGCCCTACGA
CAAGAAAAGCGGATGCTGGAAGTACCAGGCAGTGACAGCCACCCTGGAGGAGAAGAGAGAAAGC
CA
```

【図 7】

ゲノム配列 (配列 B) の配列情報の一部を示す説明図

552

.....
TATTCTTAGCTTGTGTGGCCAATTGTTGCTTATGGGGGAATGACTTTTGAAGACTTGATCTAGAGA
TGGAAATCCACAGTCTCTTTCTCATTTTCATCCAAACTGAGTCTGCTGTTTTGTGTTTTATTATAGAGCA
GTCAGGTTCCCTTCTCCCTGAAGCCAACCTAGTACCTAGGGCACTAAGATTATGTTAAGAGGCTTTTGT
GTGCTAATGTCTAATTCAAGGCTGATGGAAGTGAATTTTATCATAATATGTGAATAAAATACATTTT
TCTGAAAAAAAAGTGAGTTCTACCAAAACCAAGTGAAGGAGCCCATGATCCACCAACAGGGACTTC
TCAGCTACAAATGGGAACGTTTGTGCTCCAGCTGGGCTGCAGCTCCACCTGCAGAAATGAGGAGGAGGG
ACCACAAAGTAACAGGTGATGATCACTAACATTTCCATCATCTGCTTTTCTCTCAATGGCCAGTT
AACACAAGATGCTCTTGCACAGATGCAGAAATCTCATAAGCCATCAACTTTACCCCTGAATAGAAGTAAA
AAGGCTTTATTCATTTTCTCCCTCCCTAAATTTATTAATACCTGATAGATGICAAACACTGTTAGGT
ATGAAGATACAGTCATGAGTGAAGCATGTTCTTGGAAAGAAGACATAGCCAGCTCTCCATAGAAATGAA
ATACAGCAATAATATATATATTAATAGTTAATGGGTTTTTTTGTCTACAAAAAAACAAATTTTT
CTATCACTTAGCAAAAGTGACTAGGTCATTTTACTTTTTTGAAGTTGATTATTTGGCTAATATTATAAAT
GCCAGAGCTAAAAATAGCTGTACCTGGGGTGAAATGGAGAAGACGTGGACATAGCTTTAAAAATGGGAG
AAGCGCTTTTCCCAAGCGCTGCCGAAGATGGCGAGGTGCAGGTCTCTGGTCTCGATGGTCGAGGCCA
TCTCTGGTCCGCTGGCGGCCATCGTGGCTAAACAGGTACTGCTGGCCGGAAAGTGGTGGTGTACGC
TGCGAAGGCATCAACATTTCTGGCAATTTCTACAGAAACAAGTTGAAGTACCTGGGTTTCTCCGCAAGC
GGATGAACACCCACCTTTCCCGAGGTCCCTACCACTTCCGGCCCCCAGCCGCATCTTCTGGCGGACCG
TGCGAGGTATGCCGCCCCACAGACCAAGCGAGGCCAGGCTTCTCTGGACCGCTCAAGGTGTTTGACCG
CATCCACCGCTTACGACAAGAAAAAGCGGATGGTGTCTCTGCTCCCTCAAGGTGTGCGTCTGAAGCC
TACAAGAAAGTTTGCCTATCTGGGGCCCTGGCTCACCAGGTTGGCTGGAAAGTACCAGGCGAGTGACAGCC
ACCTTGAGGAGAGAGGAAAGAGAACCAAGATCCACTACCGGAAGAGAAGAACAGCTCATGAGGCTAC
.....

【図 8】

配列 A と配列 B との比較結果を示す説明図

.....
配列B TAAGCCATCACTTTACCTGAATAGAAAGTCTTTATTCATTTTCTCTCCCC 12313447	
配列A	
.....
配列B GGACATAGCTTTAAAAATGGGAGAGCGCTTTTCCCAAGC	802
配列A	GGCTGCCGAAGATGGCGGA 19
.....
配列B GGTGCAGGTCCTGGTGCTCGATGGTCGAGGCCATCTCTGGTCGGCTGGCGGCCATCGT 12313867	
配列A	GGTGCAAGTCCTGGTGCTCGATGGTCGAGGCCATCTCTGGTCGGCTGGCGGCCATCGT 79
.....
配列B GGCTAAACAGGTACTGCTGGGCCGGAAAGTGGTGGTCTGACGGTCGGAAGGCATCAACAT 12313927	
配列A	GGCTAAACAGGTACTGCTGGGCCGGAAAGTGGTGGTCTGACGGTCGGAAGGCATCAACAT 139
.....
配列B TTCTGGCAATTTCTACAGAAACAAGTTGAAGTACCTGGGTTTCTCCGCAAGCGGATGAA 12313987	
配列A	TTCTGGCAATTTCTACAGAAACAAGTTGAAGTACCTGGGTTTCTCCGCAAGCGGATGAA 199
.....
配列B CACCCACCTTTCCCGAGGTCCCTACCACTTCCGGGCCCCCGAGCGGCATCTTCTGGCGGA 12314047	
配列A	CACCCACCTTTCCCGAGGTCCCTACCACTTCCGGGCCCCCGAGCGGCATCTTCTGGCGGA 259
.....
配列B CCGTGCAGGTATGCCGCCCCACAAGACCAAGCGAGGCCAGGCTTCTTGACCCGCTCA 12314107	
配列A	CCGTGCAGGTATGCCGCCCCACAAGACCAAGCGAGGCCAGGCTTCTTGACCCGCTCA 319
.....
配列B AGGTGTTTGACCGCATCCACCGCCCTACGACAAGAAAAAGCGGATG 12314167	
配列A	AGGTGTTTGACCGCATCCACCGCCCTACGACAAGAAAAAGCGGATG 379
.....
配列B CTCAAGGTTGTGCTGTGAAGCCTACAAGAAAGTTTGCTATCTGGGGCGCTGGCTCAC 12314227	
配列A	
.....
配列B GAGGTTGGCTGGAAGTACCAAGCATGACAGCCACCTGGAGGAGAGAGAGAGAAA 12314287	
配列A	GCTGGAAGTACCAAGCATGACAGCCACCTGGAGGAGAGAGAGAGAAA 418
.....
配列B GCCAAGATCCACTACCGGAAGAGAAACAGCTCATGAGGCTACGGAACACAGGCCGAGAAG 12314347	
配列A	GCCA
.....

801

【図 9】

周辺遺伝子とその位置情報の一例を示す説明図

周辺 遺伝子名	位置 (bp)
ABCC13	1311191
STCH	1410233
SAMSN1	1522621
NRIP1	1998612
USP25	2767504
CXADR	4550321
BTG3	4630946
C21orf91	4826288
C21orf68	5282143
PRSS7	5306683
NCAM2	8232571
C21orf42	12337804
MRPL39	12537621
JAM2	12591245
ATP5J	12676442
GABPA	12681637
APP	12832513
CYYR1	13418319
ADAMTS1	13788256
ADAMTS5	13873414
N6AMT1	15824387
ZNF294	15880030
C21orf6	15957644
USP16	15976534
CCT8	16008212
C21orf7	16027963
BACH1	16249331
GRIK1	16487389
CLDN17	17116394
CLDN8	17164457
TIAM1	18068878
SOD1	18610012

周辺 遺伝子名	位置 (bp)
KIAA1172	18621440
HUNK	18823719
C21orf45	19219355
C21orf108	19261422
MGC14136	19343534
C21orf63	19362789
FLJ10932	19522641
C21orf59	19552201
SYNJ1	19579165
C21orf66	19706745
C21orf62	19743967
OLIG2	19976351
IFNAR2	20180297
IL10RB	20216777
IFNAR1	20275300
IFNGR2	20353293
C21orf4	20400614
C21orf55	20435928
GART	20453915
SON	20493461
ITSN1	20592797
ATP5O	20853849
MRPS6	21023940
SLC5A3	21045589
KCNE2	21314414
C21orf51	21325902
KCNE1	21399516
DSCR1	21466877
CLIC6	21658387
RUNX1	21738207
FLJ20856	21988335
C21orf18	22984669

周辺 遺伝子名	位置 (bp)
CBR1	23020125
CBR3	23085092
C21orf5	23114668
KIAA0136	23270058
CHAF1B	23335222
CLDN14	23410443
PSMD4P	23435675
SIM2	23649498
HLCS	23703739
DSCR6	23955953
DSCR5	24015109
TTC3	24032681
DSCR3	24173160
DYRK1A	24368641
KCNJ6	24574220
DSCR4	25003750
DSCR8	25070978
KCNJ15	25206098
ERG	25329397
ETS2	25755188
DSCR2	26124818
WDR9	26135251
HMG14	26290570
WRB	26326951
SH3BGR	26392496
B3GALT5	26602941
DSCAM	27084862
BACE2	28110098
C21orf11	28259086
MX2	28312529
MX1	28368500
TMPRSS2	28406900

【図 10】

未知遺伝子の配列 A と周辺遺伝子の距離を算出した結果を示す説明図

周辺 遺伝子名	未知 遺伝子 からの 距離 (bp)	周辺 遺伝子名	未知 遺伝子 からの 距離 (bp)	周辺 遺伝子名	未知 遺伝子 からの 距離 (bp)
C21orf42	24015	SYNJ1	7265376	C21orf5	10800879
MRPL39	223832	C21orf66	7392956	STCH	10903556
JAM2	277456	C21orf62	7430178	KIAA0136	10956269
ATP5J	362653	C21orf91	7487501	ABCC13	11002598
GABPA	367848	OLIG2	7662562	CHAF1B	11021433
APP	518724	BTG3	7682843	CLDN14	11096654
CYYR1	1104530	CXADR	7763468	PSMD4P	11121886
ADAMTS1	1474467	IFNAR2	7866508	SIM2	11335709
ADAMTS5	1559625	IL10RB	7902988	HLCS	11389950
N6AMT1	3510598	IFNAR1	7961511	DSCR6	11642164
ZNF294	3566241	IFNGR2	8039504	DSCR5	11701320
C21orf6	3643855	C21orf4	8086825	TTC3	11718892
USP16	3662745	C21orf55	8122139	DSCR3	11859371
CCT8	3694423	GART	8140126	DYRK1A	12054852
C21orf7	3714174	SON	8179672	KCNJ6	12260431
BACH1	3935542	ITSN1	8279008	DSCR4	12689961
NCAM2	4081218	ATP5O	8540060	DSCR8	12757189
GRIK1	4173600	MRPS6	8710151	KCNJ15	12892309
CLDN17	4802605	SLC5A3	8731800	ERG	13015608
CLDN8	4850668	KCNE2	9000625	ETS2	13441399
TIAM1	5755089	C21orf51	9012113	DSCR2	13811029
SOD1	6296223	KCNE1	9085727	WDR9	13821462
KIAA1172	6307651	DSCR1	9153088	HMG14	13976781
HUNK	6509930	CLIC6	9344598	WRB	14013162
C21orf45	6905566	RUNX1	9424418	SH3BGR	14078707
C21orf108	6947633	USP25	9546285	B3GALT5	14289152
PRSS7	7007106	FLJ20856	9674546	DSCAM	14771073
MGC14136	7029745	NRIP1	10315177	BACE2	15796309
C21orf68	7031646	C21orf18	10670880	C21orf11	15945297
C21orf63	7049000	CBR1	10706336	MX2	15998740
FLJ10932	7208852	CBR3	10771303	MX1	16054711
C21orf59	7238412	SAMSN1	10791168	TMPRSS2	16093111

【図 11】

発現プロファイルデータベースの内容の一例を示す説明図

1101	ID	Hs.8230	...	1102
	TITLE	a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 1		
	GENE	ADAMTS1		
	CYTOBAND	21q21.2		
	LOCUSLINK	9510		
	EXPRESS	uterus (pregnant);placenta;multiple sclerosis lesions;fibroblast;bone;thyroid gland; fetal liver/spleen;infant brain;fetal heart;fetal lung;melanocyte;gall bladder; pancreatic islets;ovary;lung;fetal cochlea;breast;embryo (12 week);heart		
	GNM_TERMINUS	T		
	CHROMOSOME	21		
	STS	ACC=G32278 NAME=A005R14 UNISTS=116876		
	STS	ACC=- NAME=A005R14 UNISTS=61779		
	STS	ACC=G06805 NAME=D21S1964 UNISTS=63269		
	STS	ACC=- NAME=sts-R76276 UNISTS=32124		
	STS	ACC=G14513 NAME=SHGC-11277 UNISTS=61020		
	STS	ACC=G34580 NAME=RH43649 UNISTS=61021		
	TXMAP	D21S265-D21S260; MARKER=A005R14; RHPANEL=G84		
	TXMAP	D21S260-D21S261; MARKER=sts-R76276; RHPANEL=G84		
	TXMAP	D21S265-D21S260; MARKER=SHGC-11277; RHPANEL=G3		
	PROTSIM	ORG=Caenorhabditis elegans; PROTG1=7499847; PROTID=pir:T21371; PCT=44; ALN=705		
	PROTSIM	ORG=Homo sapiens; PROTG1=11360273; PROTID=pir:T47158; PCT=99; ALN=966		
	PROTSIM	ORG=Mus musculus; PROTG1=7513658; PROTID=pir:T00017; PCT=82; ALN=966		
	PROTSIM	ORG=Rattus norvegicus; PROTG1=11131014; PROTID=sp:Q9WUQ1; PCT=82; ALN=966		
	SCOUNT	227		
	SEQUENCE	ACC=NM_006988; NID=g11038653; PID=g11038654		
	SEQUENCE	ACC=AF207664; NID=g6685071; PID=g6685072		

【図 12】

周辺遺伝子の発現部位情報の一例を示す説明図

周辺遺伝子名	発現部位
C21orf42	brain; germ cell; lymphoma, cell line; pool; testis
MRPL39	Stomach; adenocarcinoma; adrenal gland; blood; brain; breast; carcinoma, cell line; cervical carcinoma cell line; chondrosarcoma; choriocarcinoma; colon; denis_drash; duodenal adenocarcinoma, cell line; fetal eyes; fetal eyes, lens, eye anterior segment, optic nerve, retina, retina foveal and macular, rpe and choroid; foreskin; from chronic myelogenous leukemia; glioblastoma; glioblastoma with egr amplification; head_neck; head_normal; heart; hippocampus, cell line; hypothalamus, cell line; kidney; leiomyosarcoma; liver; lung; lung_normal; melanoma (mewo cell line); melanotic melanoma; neuroblastoma; ovary; parathyroid; placenta; pool; pooled; primitive neuroectoderm; prostate; retina foveal and macular; rpe and choroid; subchondral bone; testis; thymus, pooled; tonsil; uterus; whole embryo
JAM2	brain; carcinoid; ear; fetal eye; germ cell; glioblastoma; glioblastoma with egr amplification; heart; hypothalamus, cell line; kidney; lung; lung_normal; nervous_normal; nervous_tumor; neuroblastoma cells; neuroblastoma, cell line; ovary (pool of 3); pancreas; placenta; pool; pooled; pooled brain, lung, testis; prostate; retina foveal and macular; rpe and choroid; subchondral bone; testis; testis, cell line; testis_normal; three pooled meningiomas; uterus
ATP5J	Stomach; adenocarcinoma; adrenal cortex carcinoma, cell line; adrenal cortico adenoma for cushing's syndrome; adrenal gland; aorta; ascites; bladder_tumor; blood; bone; bone marrow; brain; breast; carcinoid; carcinoma, cell line; colon; colon_normal; cord blood; correspond
GABPA	Stomach; adenocarcinoma; amygdala; blood; brain; breast_normal; carcinoid; cns; colon; duodenal adenocarcinoma, cell line; embryonal carcinoma; embryonal carcinoma, cell line; foreskin; germ cell; head_neck; heart; hippocampus, cell line; kidney; leiomyosarcoma; lung; lymphoma, cell line; mammary adenocarcinoma, cell line; medulla; neuroblastoma cells; ovary; placenta; pool; retinoblastoma; small intestine; testis; uterus
APP	2 pooled high-grade transitional cell tumors; 2 pooled wilms' tumors, one primary and one metastatic to brain; Fetal brain; Prostate; Stomach; adenocarcinoma; adenocarcinoma cell line; adenocarcinoma,
CYR1	2 pooled wilms' tumors, one primary and one metastatic to brain; adrenal gland; brain; breast; breast_normal; carcinoid; chondrosarcoma; cns; follicular carcinoma; from acute myelogenous leukemia; hepatocellular carcinoma; insulinoma; kidney; lung; metastatic chondrosarcoma; ovary (pool of 3); parathyroid; pool; pooled; prostate; testis; testis_normal; uterus; whole embryo

【図 13】

周辺遺伝子の発現部位情報の別の一例を示す説明図

周辺遺伝子名	発現部位
C21orf42	brain; germ cell; lymphoma, cell line; pool; testis
MRPL39	Stomach; adenocarcinoma; adrenal gland; blood; breast; carcinoma, cell line; cervical carcinoma cell line; chondrosarcoma; choriocarcinoma; colon; denis_drash; duodenal adenocarcinoma, cell line; fetal eyes; fetal eyes; lens, eye anterior segment, optic nerve, retina, retina foveal and macular, rpe and choroid; foreskin; from chronic myelogenous leukemia; glioblastoma; glioblastoma with egfr amplification; head_neck; head_normal; heart; hippocampus, cell line; hypothalamus, cell line; kidney; leiomyosarcoma; liver; lung; lung_normal; melanoma (mewo cell line); melanotic melanoma; neuroblastoma; ovary; parathyroid; placenta; pooled; primitive neuroectoderm; prostate; retina foveal and macular; rhabdomyosarcoma; thymus, pooled; tonsil; uterus; whole embryo
JAM2	carcinoid; ear; fetal eye; nervous_normal; nervous_tumor; neuroblastoma cells; neuroblastoma, cell line; ovary (pool of 3); pancreas; pooled brain, lung, testis; rpe and choroid; subchondral bone; testis, cell line; testis_normal; three pooled meningiomas
ATP5J	adrenal cortex carcinoma, cell line; adrenal cortico adenoma for cushing's syndrome; aorta; ascites; bladder_tumor; bone; bone marrow; colon_normal; cord blood; correspond
GABPA	amygdala; breast_normal; cns; embryonal carcinoma; embryonal carcinoma, cell line; mammary adenocarcinoma, cell line; medulla; retinoblastoma; small intestine
APP	2 pooled high-grade transitional cell tumors; 2 pooled wilms' tumors, one primary and one metastatic to brain; Fetal brain; Prostate; adenocarcinoma cell line; adenocarcinoma, cell line; adipose; adrena
CYR1	follicular carcinoma; from acute myelogenous leukemia; hepatocellular carcinoma; insulinoma; metastatic chondrosarcoma
ADAMTS1	amion_normal; colonic mucosa from 3 patients with crohn's disease; duodenal adenocarcinom
ADAMTS5	muscle

【書類名】 要約書

【要約】

【課題】 未知遺伝子の発現部位予測・機能解析を迅速かつ効率的におこなうことが可能な遺伝子発現部位予測支援方法、遺伝子発現部位予測支援プログラム、遺伝子発現部位予測支援装置を提供すること。

【解決手段】 未知遺伝子の配列を入力し（ステップS 5 0 1）、対象遺伝子のゲノム配列へのm a p p i n gをおこない（ステップS 5 0 2）、ゲノム配列上における、発現部位が特定されている周辺遺伝子が存在する位置と未知の遺伝子が存在する位置との距離を算出し（ステップS 5 0 3）、算出された距離に基づいて、前記未知の遺伝子の発現部位を抽出する（ステップS 5 0 4）。より具体的には、算出された距離が短い一つまたは複数の周辺遺伝子の発現部位を前記未知の遺伝子の発現部位として特定する。

【選択図】 図 5

特願 2 0 0 2 - 2 2 4 0 7 5

出 願 人 履 歷 情 報

識別番号

[0 0 0 0 0 5 2 2 3]

1. 変更年月日

1 9 9 0 年 8 月 2 4 日

[変更理由]

新規登録

住 所

神奈川県川崎市中原区上小田中 1 0 1 5 番地

氏 名

富士通株式会社

2. 変更年月日

1 9 9 6 年 3 月 2 6 日

[変更理由]

住所変更

住 所

神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号

氏 名

富士通株式会社